DOCUMENT RESUME

ED 368 097 EC 302 842

AUTHOR Johnson, Nancy E.
TITLE Use of the WISC-R with Disadvantaged Gifted Children: Current Practice, Limitations, and Ethical Concerns.
PUB DATE 94
CONTRACT R206A00569
NOTE 36p.; In: Saccuzzo, Dennis P.; And Others. Identifying Underrepresented Disadvantaged Gifted and Talented Children: A Multifaceted Approach. (Volumes 1 and 2); see EC 302 840.
PUB TYPE Reports - Research/Technical (143) -- Information Analyses (070)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Aptitude Tests; Comparative Analysis; Cultural Differences; *Culture Fair Tests; Demography; Disadvantaged Youth; Educational Diagnosis; Elementary Education; Elementary School Students; Ethnic Groups; *Gifted; Intelligence Tests; *Minority Group Children; *Psychometrics; Student Characteristics; *Talent Identification; *Test Bias; Testing Problems; Test Validity
IDENTIFIERS San Diego Unified School District CA; *Wechsler Intelligence Scale for Children (Revised)

ABSTRACT

This pair of studies investigated the psychometric adequacy of the Wechsler Intelligence Scale for Children - Revised (WISC-R) with a population of 8,396 potentially gifted elementary children of diverse ethnic and cultural backgrounds as well as diverse emotional and social environments. Study I showed subjects' profiles differed strikingly from those of the WISC-R standardization sample. It also found subjects differed strikingly across ethic groups and across levels of risk in the shape of the verbal IQ-performance IQ difference distribution but not in absolute size of this difference. Skewness of profiles of children with identified risk was in favor of performance IQ relative to those without risk. Study II attempted to identify either a single model or individual models to select equally accurately from five ethnic groups (African-American, Asian, Caucasian, Filipino, and Hispanic). No single model or combination of individual models was found. Implications for this use of the WISC-R in diverse gifted populations whose characteristics differ from those of the standardization sample are discussed in the context of the professional ethics of responsible test use. Forms used in the San Diego gifted identification process are attached. (PB)

CHAPTER 2

Use of the WISC-R with Disadvantaged Gifted Children: Current Practice, Limitations, and Ethical Concerns

Nancy E. Johnson
San Diego State University/University of California San Diego
Joint Doctoral Program in Clinical Psychology

This research was funded in part by Grant R206A00569, U.S. Department of Education, Jacob Javits Gifted and Talented Discretionary Grant Program.

The author expresses appreciation to her mentor, Dennis P. Saccuzzo, San Diego State University, to the San Diego Unified City Schools, to Gifted and Talented Education (GATE) Administrator David P. Hermanson, and to the following school psychologists: Will Boggess, Marcia Dijiosia, Eva Jarosz, Dimaris Michalek, Lorraine Rouse, Ben Sy, and Daniel Williams.

Correspondence should be addressed to Nancy E. Johnson, San Diego State University, 6363 Alvarado Court, Suite 103, San Diego, California 92120-4913 (Telephone: 619-594-2845 /FAX: 619-594-6780 / e-mail: njohnson@psychology.sdsu.edu).

ABSTRACT

The Wechsler Intelligence Scale for Children-Revised (WISC-R) is the most widely used individual instrument for inclusion or exclusion of children into programs for the gifted in the United States. The present study investigated the psychometric adequacy of this use of the WISC–R in a population of 8396 potentially gifted elementary grade children of diverse ethnic and cultural backgrounds as well as diverse emotional and social environments. Study I included analyses of VIQ-PIQ base rates in 5796 children who achieved Full Scale IQ (FSIQ) scores of 130 or above, plus comparisons of similarities and differences in subsamples divided on ethnic background, on level of risk identified in the child's home environment, and on the extremes of achievement (measured by a standardized achievement test). In contrast to findings from the WISC-R standardization sample, children in this study differed strikingly, across ethnic groups and across levels of risk, in shape of the VIQ-PIQ difference distribution but not in absolute size of the VIQ-PIQ difference. The frequency distributions for African-Americans and Caucasians were skewed in favor of VIQ; for Filipinos, in favor of PIQ. Those of Asians and Hispanics closely resembled normal distributions. Skewness for children with identified risk was in favor of PIQ relative to those without risk. The importance of clinical versus statistical significance in decision-making was discussed, with particular attention to what constitutes a 'rare' VIQ-PIQ difference in gifted children. Study II attempted, through multivariate modeling, to identify either a single model or individual models, using subtests of the WISC-R, that would select equally accurately from five ethnic groups (African-American, Asian, Caucasian, Filipino, and Hispanic). No single model or combination of individual models was found to select equally from each of the ethnic backgrounds in a proportionately balanced random subsample of 1438. Implications for this use of the WISC-R in diverse gifted populations whose characteristics differ from those of the standardization sample were discussed, in light of the professional ethics of responsible test use.

I. Introduction

Currently, identification of giftedness in school age children is undertaken nationwide with the aim of providing special educational services for those with special gifts and talents. Historically, the use of tests to identify individuals with special talent has been recorded as early as 2200 BC in China (DuBois, 1970). In 1869 Galton first addressed the concept of genius in the psychological literature. In 1925 Terman began the first major study in which giftedness was operationally defined in terms of performance on standardized IQ tests. Since these landmark contributions, conflict and controversy have abounded in the educational and psychological literature on giftedness. Disagreements continue over the definition of giftedness per se, its measurement by the use of IQ and achievement tests, and its nurturance by special instructional programs. This work will focus on one aspect of giftedness: ethical use of tests in the selection of children from diverse backgrounds for early inclusion in special programs for the gifted and talented.

Identification and inclusion of gifted children from varied cultural and linguistic backgrounds into gifted and talented programs at an early age is vital. As Horowitz and O'Brien (1986) note, "there is no way to measure the loss when individuals capable of functioning considerably above the normal level do not contribute as much to society as their capabilities will allow" (p. 1147). The summary of findings in an evaluation of the Gifted and Talented Education (GATE) program in San Diego in the academic year 1989-1990 (Millett, 1990) included the information that "GATE students outperformed gifted students who are not participating in the program at every grade level" (p. 9). Given the demonstrated benefits of programs for gifted children, educators face the challenge of early identification of children with the highest potential for inclusion in enrichment programs. This problem becomes more critical in light of mandates that educational programs strive to guarantee equal access and yet operate within a framework of increasingly restrictive educational budgets.

Problems in the Definition of Giftedness

Currently in this country most efforts to identify giftedness in children utilize a definition based on intelligence, measured by some form of standardized group or individual IQ test. The trend began to be formalized in 1971, when the first definition of gifted and talented children was proposed on a national level (Pub. L. 91-230, § 806):

> Gifted and talented children are those identified by professionally qualified persons who by virtue of outstanding abilities, are capable of high performance. These are children who require differentiated educational programs and/or services beyond those normally provided by the regular school programs in order to realize their contribution to self and society.
>
> Children capable of high performance include those with demonstrated achievement and/or potential ability in any of the following areas, singly or in combinations: (1) general intellectual ability, (2) specific academic aptitude, (3) creative or productive thinking, (4) leadership ability, (5) visual and performing arts, (6) psychomotor ability.

Seven years later, 42 states had either enacted laws or formulated guidelines for the definition of giftedness. In all 42 states, including California and New York, general intellectual ability was specified (Fox, 1981).

The issue of the definition of the nature of "intelligence," thought to underlie intellectual ability and academic performance, has probably been debated as much as any other in the history of the psychological literature. Binet (in Terman, 1916) defined intelligence as "the capacity to make adaptations for the purpose of attaining a desired end" (p. 45). Spearman (1923) wrote that intellect involves "educing either relations or correlates" (p. 300), and proposed a two-factor theory; *g* was defined as an underlying general mental energy, whereas s represented one or more specific factors. Wechsler (1958) espoused the definition: "the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment" (p. 7). However, Thorndike (1927) theorized that intelligence involves interconnected but distinct abilities and so advocated a multifactor approach.

Guilford (1967) developed a multifactor theory of intelligence based on three dimensions—the operations involved in information processing, the contents, and the products. In contrast, Vernon's (1950) was a hierarchical theory of intelligence based on the hypothesis that *g* is at the highest level of the hierarchy and represents the broadest aspect.

More recently, Sternberg (1986) developed a theory that divides intelligence into three dimensions. Gardner (1983), on the other hand, suggested that there are several distinct and relatively separate competencies, which he described as multiple intelligences. The debate continues, with some theorists espousing models based on an underlying basic mental capacity and others favoring a set of distinct and relatively discrete mental abilities.

Issues in the Assessment of Intellectual Giftedness

In acknowledging that there are many definitions of what constitutes intelligence, we must also acknowledge that there are many tests that purport to measure it. At the present time, however, the single instrument most frequently used for identification of giftedness in children in the United States is the Wechsler Intelligence Scale for Children - Revised (WISC-R) (Klausmeier, Mishra, & Maker, 1987). The WISC-R has been widely acknowledged to have excellent reliability and concurrent validity (Sattler, 1988). The current study examined the characteristics, efficacy, and fairness of this particular use of the WISC-R in one large school district (San Diego City Schools) over a seven-year span of time.

The San Diego City School District is among the most culturally diverse in the nation. The 1991-92 student population of 123,503 included 35.4% Caucasian, 28.8% Hispanic, 16.3% African American, and 8.1% Filipino children. The remaining 11.4% consisted of Indochinese, Asian, Pacific Islander, and Native American students. Programs for gifted and talented students, begun in the district in the 1940's, have demonstrated an on-going commitment to achieving equal access for individuals of all ethnic backgrounds through the use of selection instruments more likely to identify giftedness in culturally and linguistically different students. Despite these attempts, the non-Caucasian student population in gifted programs was 36.3% in 1989-90, as opposed to 61% in the school district as a whole. Underrepresented groups included Hispanics and African Americans; overrepresented were Asian, Filipino, and non-Hispanic white students (Millett, 1990). Richert (1987) cited figures published by the U.S. Department of Education's Office of Civil Rights revealing that groups such as Hispanics and African-Americans are underrepresented by as much as 70% in gifted programs throughout this nation. Thus the underselection of these two groups in San Diego reflects a nationwide problem. The National Report on Identification for Gifted and Talented Youth (Richert, Alvino, & McDonnel, 1982) noted problems with traditional selection procedures. Indeed, today most authorities believe that especially for disadvantaged groups traditional standardized tests should not be the sole or even the primary measure of giftedness (Fox, 1981; Garcia, 1981; Horowitz & O'Brien, 1986; McKenzie, 1986; Meeker & Meeker, 1973; Renzulli, 1978; Sternberg, 1981).

The American Educational Research Association, The American Psychological Association, and the National Council on Measurement in Education take the position that "In elementary or secondary education, a decision or characterization that will have a major impact on a test taker should not automatically be made on the basis of a single test score." (Standard 8.2, Standards for Educational and Psychological Testing, 1985). Although many authorities do recommend the use of multiple identification procedures such as IQ, achievement, and behavioral data in the identification of giftedness, in practice much emphasis is commonly placed on a single measure of achievement or of overall intelligence (Alvino, McDonnel, & Richert, 1981). Zigler and Farber (1985) stated that a specific defined level of IQ (such as a score two standard deviations above the mean) is the most adequate index of giftedness. Pegnato and Birch (1959), Clark (1979), and Hagen (1980) recommended use of an individually administered IQ test as the best and the quickest way to find most gifted children. Sattler (1988) concluded that "the single best method available for the identification of children with superior cognitive abilities is a standardized, individually administered test of intelligence..." (p. 671), but went on to note that among those who are difficult to identify as gifted are children who are culturally different, especially since they may not show superior oral language skills. Indeed, as was so well expressed by the Standards for Educational and Psychological Testing (1985), "A child from one culture

who is evaluated with morés appropriate to another culture may be considered taciturn, withdrawn, or of low mental ability."

Methodological Issues in WISC-R Testing of Ethnic Groups

Use of standardized intelligence test summary scores without ethnic, cultural, gender, economic, and other considerations is based on a uniformity assumption: that all students, all testers, and all situations are homogeneous. The fallacies inherent in this assumption in the use of standardized tests have been repeatedly noted (Guertin, Ladd, Frank, Rabin, & Hiesler, 1971; Lewandowski & Saccuzzo, 1976). Unfortunately, most standardized tests have only a single set of norms that have not been corrected for the demographic characteristics of the individual. The WISC-R, for example, yields scores corrected only for chronologic age. It has long been recognized that the influence of demographic variables in tests of brain function is apparent for individuals (Finlayson, Johnson, & Reitan, 1977; Reitan, 1955). For example, recent cross-sectional studies of the Wechsler tests for adults indicate that a single set of norms cannot be used for individuals at different age and education levels (Heaton, Grant, & Matthews, 1986).

Further, the use of a single summary score may mask differences in the pattern of strengths across ethnic backgrounds and gender in gifted children. Lesser, Fifer, and Clark (1965) reported results of a comparison of mental abilities in seven and eight year old first grade children from four ethnic groups and two socioeconomic levels in New York. Individuals of African-American, Chinese, Jewish, and Puerto Rican background were compared on four basic dimensions of mental ability using a modified version of the Hunter College Aptitude Scales for gifted children. The children were found to differ in pattern of mental abilities across ethnic background but not across socioeconomic status. Lesser et al. proposed that identification of the pattern of relative strengths and weaknesses of children from varied cultural backgrounds was a vital prerequisite to decisions about education in general and curriculum in particular.

Methodological Issues in Quasi-Experimental Assessment Studies

In reviewing the literature on the use of tests with different ethnic groups, several methodoligical issues become apparent. Some are inherent in the nature of quasi-experimental and archival design (e.g., the impossibility of random assignment to groups on key factors such as ethnic background or socioeconomic status), and limit the generalizability and applicability of the studies. Others result from a failure to control for moderator variables such as socioeconomic status and acculturation, or from a failure to use multiple methods within the same study. Several of these points will be illustrated in the following examination of studies subsequent to Lesser, et al.

Since publication of Lesser et al.'s findings, numerous investigators have made attempts to confirm differences in pattern of mental abilities across ethnic groups (e.g., Flaugher & Rock, 1972; Hennessy & Merrifield, 1976; Sitkei & Meyers, 1969). None of the subsequent studies have used the same tests, the same ethnic groups, or even children of the same ages. Most did not control for level of ability, and no single study looked at all of these confounds systematically. Despite these flaws, there has been a tendency among reviewers (e. g., Sattler; 1988) to characterize these attempts as "failure to replicate" the findings of Lesser et al.. Indeed, Sattler cited one study of 4 year olds (Sitkei & Meyers, 1969), one of junior high school students (Flaugher & Rock, 1972), and one of high school seniors accepted for admission to a major university (Hennessy & Merrifield, 1976) as evidence of failure to replicate. These studies were factor analytic in nature. Whereas the original work by Lesser et al. was based on an analysis of covariance method of comparing mean scores on tests across groups, the studies cited by Sattler, as well as elsewhere in the literature, compared factor structure of a given test across groups, and some included a comparison of factor means across the groups.

Several issues in quasi-experimental design and methodology become apparent when such studies are compared as "replications":

1. It is not valid to compare results of studies with populations of preschool children, elementary school children, junior high school students, and high school seniors. The increased exposure to

environments outside the cultural environment of the home as the child progresses through school is, for example, an enormous confound and provides a valid alternative hypothesis for the different findings.

2. Different assessment batteries can produce different results. Sitkei and Meyers (1969) used an extensive battery that included the Peabody Picture Vocabulary Test, which, as the authors acknowledged, is much less highly verbally loaded than the measures used by Lesser et al.. In fact, Sitkei and Meyers offered this lower verbal demand as one possible alternative hypothesis for the difference in their findings from those of Lesser et al..

3. Results at one level of intelligence do not necessarily generalize to others. Lesser et al. studied children matched on the basis of social class, gender, and ethnic membership; each of those matching variables has been correlated with differences in performance on tests of mental ability. Hennessy and Merrifield's (1976) subject pool was restricted to high school seniors who had been accepted for admission to universities in the fall. It seems unlikely that the two populations were comparable in their basic levels of mental ability, although Hennessy and Merrifield were careful to partial out the effects of socioeconomic status.

4. An analysis of covariance, directly comparing group means on subtest scores, provides different information than a factor analytic comparison, including a comparison of the factor means. Factor analysis is a data reduction technique for mathematically analyzing the intercorrelations between members of a set of variables and thus deducing a smaller set of factors. Those factors are assumed to account for the intercorrelations seen in the directly measurable original variables. The factors are arbitrarily named and interpreted (hopefully based on a theoretical model of the construct being studied); comparing factor means is not the same thing as comparing observable test score mean differences. A test could measure the same underlying mental abilities in four groups and yet produce a very different pattern of strengths and weaknesses in subtest performance across those four groups. In other words, it is possible that the groups show the same pattern of intercorrelations between subtests, but differ in the level of their original mean scores on subtests that critically load on a given factor. Group A could have consistently lower scores than Group B on all measures loading on Factor 1, and still show the same overall pattern of intercorrelations between those tests.

5. A difference in group means does not imply that most individuals in a group will have scores that fall in the direction of the observed group mean. Methodological rigor demands analysis of not only group means, but also individual data in conjunction with the group data. As Guertin, Frank, and Rabin (1956) point out: "One methodological shortcoming is the failure to distinguish between a mean diagnostic group profile and modal patterns of homogeneous subjects ... Only modal patterns are appropriate for diagnostic purposes" (p.239). For example, in a study of the WISC as a clinical diagnostic tool, Saccuzzo and Lewandowski (1976) found group differences on one subtest (Picture Arrangement) that would indicate that a preponderance of the individual scores could be expected to fall above the mean in the higher group. When individual cases were examined, however, it was found that less than half of the cases actually were above the mean, and there were no consistent tendencies on this subtest. Therefore, the subtest could not be used as a clinical indicator. In another example, these investigators found no group differences between the races in terms of WISC responses on a number of Wechsler's hypotheses regarding acting-out adolescents. On post-hoc analyses, however, there were clear differences between white males and black females that were masked by the overall means. If the issue is one of fairness of selection criteria, then individual scores must be examined in light of group means.

Again, the basic issue in the use of any test to select students for special programs is one of test use; fairness demands that the test be used in a way that will select equally from various groups, rather than invariably favoring (or disfavoring) members of one group over another. It is certainly possible to design a test that appears to measure the same underlying constructs across groups, and still find that the test differentially selects members of one group over another because of the way it is being used. That may be the case with the common practice of using the WISC-R to identify intellectual giftedness in children.

Preliminary studies with a San Diego gifted population using a group measure of intelligence, the Developing Cognitive Abilities Test, indicated that predictors of giftedness depend on ethnic background (Saccuzzo, Hermanson, Dorne, Johnson, & Shamieh, 1990). For African-Americans, the quantitative score proved most predictive, while for Hispanics the spatial and total scores were most predictive of selection for gifted programs. Total scores alone were most predictive for only the Caucasians and Filipinos, who were overrepresented in the gifted and talented program. These findings suggested that giftedness may be expressed in unique patterns of abilities not best measured by a summary IQ score. Although the study was not (and was not intended to be) a replication of Lesser et al.'s work, the results did add weight to the idea that identifiable differences exist in the way giftedness is expressed across ethnic and cultural backgrounds. In further support of this hypothesis, a summary of academic performance of all students in gifted programs in San Diego City Schools indicated that Hispanic and African American students at all grade levels generally fall below other groups (and below the 90th percentile) only in reading and language (Millett, 1990). Analysis of VIQ - PIQ discrepancies in a random subset of this population also revealed differences that varied across ethnic background and as a function of the size of the discrepancy (Saccuzzo, Johnson, & Russell, 1992).

Given that the WISC-R is one of the single most widely used instruments for the identification of giftedness in the United States, and given the problem of underselection of certain ethnic groups, the goal of this study was to examine the feasibility of using the WISC-R in any way to select a balanced population of gifted children,since it would appear that a summary IQ score will not do so. The present work began with an analysis (Study I) of the WISC-R Verbal, Performance, and Full Scale scores of children who achieved a Full Scale IQ score at least two standard deviations above the mean (FSIQ $\geq$ 130). Children were compared and contrasted in terms of basic demographic factors such as ethnicity and gender, as well as on environmental factors thought to place them at risk for limited expression of their full potential (e.g., economic, language, and emotional factors). Verbal-Performance differences were examined in a study of base rates for the entire sample of intellectually gifted children as well as for subsamples defined on the basis of ethnic background, areas of risk, and documented low or high school achievement test scores. In spite of excellent discussions by Kaufman (1976) and Matarazzo and Herman (1984) on the difference between statistical and clinical significance, little has been documented about the relative rarity of specific VIQ-PIQ discrepancies in different populations of children. Finding that a child has a statistically significant VIQ-PIQ difference tells the clinician or educator nothing more than that the difference is probably real and not due to chance: it does not address the issue of the rarity of that difference in a given population or of its real world significance, nor does it address the likelihood that such a VIQ-PIQ difference is associated with low achievement. Only by studying actual occurrence in a population can we address such issues. Kaufman (1976) noted no differences in base rates across ethnic backgrounds in children with IQ values of at least 120 in the standardization sample. Two serious problems with that finding are that Kaufman did not take into account the direction of the difference (only the size), and that there were almost certainly not enough non-Caucasian children in the sample at those IQ levels to have found a difference even if it existed: the total number at that IQ level was 213. The present study was undertaken to provide accurate base rates for a large, culturally diverse sample of gifted children, with the hope that more evidence could be provided to dispel invalid uniformity assumptions and to shed light on this gifted population.

Study II examined the feasibility of deriving a single set of criteria from the WISC-R to select a proportionately representative, ethnically diverse sample of children for inclusion in programs for the gifted by exploring two alternative hypotheses: (1) there exists a single pattern of WISC-R subtest scores that predicts giftedness equally across gender and ethnic background; or (2) there is a unique pattern of cognitive strengths and thus different predictors of giftedness for each group. Ethnic backgrounds represented included African-American, Asian, Caucasian, Filipino, and Hispanic.

General Considerations in the Use of Tests for Giftedness

Again, the basic issue is one of competent test use. Despite ongoing discussion, acknowledgment of the limitations of IQ tests, and exhortations to use these tests in an informed manner (Borland, 1986; Kaufman & Harrison, 1986; Robinson & Chamrad, 1986; Sternberg, 1982), no single

study to date in the educational or psychological literature has directly and adequately addressed the issue of fairness of the use of this test in an ethnically diverse population of potentially gifted children. Goals of this study included possible explication of a more fair and adequate use of the WISC-R in identification of giftedness, a discussion of selection bias that results from its use, and further understanding of the limits as well as the full potential of the WISC-R in the selection of students from diverse backgrounds for gifted programs in schools.

Benefits to be gained from improved methods of selection are substantial. If we are to increase the number of underrepresented minorities in the professions, as morally and legally mandated, it is vital to identify and encourage those individuals as early as possible. What are the consequences if we continue to fail in this endeavor? They are perhaps best summed by D. D'Souza (1991) in his description of the experience of one university noted for its aggressive affirmative action policy:

> ...the academic difficulties encountered by affirmative action students who find it impossible to compete effectively with other, better-prepared students, are reflected in Berkeley's extremely high dropout rate for Hispanic and black undergraduates. Whites and Asians graduate from Berkeley at about the same rate: 65-75 percent. That is to say that 25-35 percent drop out before graduation. Hispanics graduate at under 50 percent. More than half drop out. Blacks graduate at under 40 percent. More than 60 percent drop out.
>
> ...Berkeley does not release the number of blacks and Hispanics admitted on affirmative action who drop out, but these data are contained in a confidential internal report which tracks freshmen enrolled in 1982. By 1987, five years later, only 18 percent of blacks admitted on affirmative action had graduated from Berkeley; blacks admitted in the regular program graduated at a 42 percent rate. Similarly, only 22 percent of affirmative action Hispanics finished in five years, compared with 55 percent for other Hispanics. The most recent figures suggest that approximately 30 percent of black and Hispanic students drop out before the end of their freshman year; in the words of the report, they seem to stay "only long enough to enhance the admissions statistics." (p. 39)

I would propose that the key phrase is "better prepared" students and suggest that such preparation must begin as early in elementary education as possible.

Inclusion of more equitable proportions of high risk children in gifted programs is a goal much sought in education. A unique opportunity exists in San Diego to study selection procedures for gifted and talented programs: a large, ethnically diverse metropolitan population plus a school district that continues to demonstrate its commitment to identification of underrepresented and disadvantaged students.

II. Methods

Two studies were completed: I) an analysis of the base rates of VIQ-PIQ differences in a population of intellectually gifted children, defined as those who achieve a Full Scale IQ score at least two standard deviations above the mean, and II) an examination of the use of the WISC-R to select a proportionately representative and ethnically diverse sample of gifted children from the population of children identified as potentially gifted.

Subjects

Each child in this study was identified as potentially gifted based on achievement test data, teacher evaluation (Appendix A) and recommendation, and a social case study analysis (Appendices B and C). The social case study analysis included an assessment of 6 areas of potential risk for achievement and expression of full potential: 1) cultural, 2) economic, 3) emotional, 4) environmental, 5) health, and 6) language. Cultural risk included cultural values and beliefs that differ from the those of the dominant culture, or limited experience in the dominant culture. Economic risk included parental unemployment or household income low enough to qualify the child for the free lunch program. Emotional risk

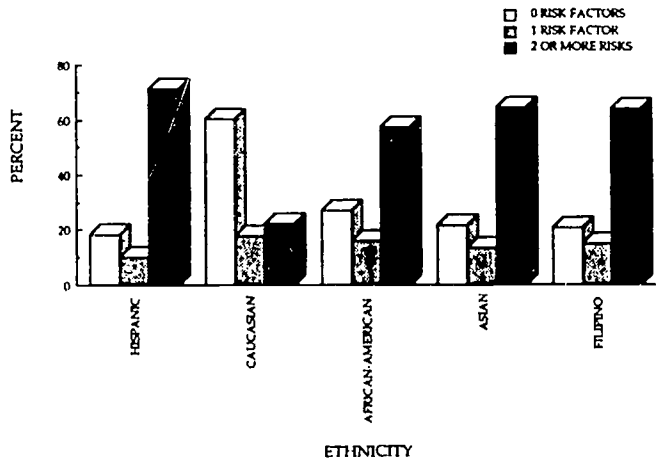
encompassed such factors as death of a parent, child abuse, major psychiatric illness in the home, or extended absence of a parent due to military service. Environmental risk included transiency (three or more school moves) and excessive absences from school due to home responsibilities such as child care responsibility or working to help support the family. Health factors included vision, speech, or hearing deficits requiring designated instructional service, motor problems requiring adaptive physical education, or diseases such as asthma. Children at risk due to language included those for whom English is a second language and those not fluent in English. For the purposes of the current project, each child was assigned a value for level of risk: 0 if no identified risk, 1 if risk was identified in one and only one of the areas described above, and >1 if more than one area of risk was identified for that child.

Ethnic background was determined by self-report, based on an information questionnaire completed by parents at the time of their child's enrollment in the school district. Problems are inherent in such self-report, including the resultant heterogeneity of each group. For example, the child of one Caucasian and one Hispanic parent may be reported to be Caucasian or to be Hispanic, depending on societal factors that enter into the parents' decision to report: Being considered Caucasian might seem to confer some obvious dominant culture benefits, but being designated Hispanic might open opportunities for scholarship or for special tutorial programs in a given school.

Ethnic categories designated by this district are broad and in themselves create heterogenious groups: 'Hispanic' includes those from Mexico, Cental and South America, Puerto Rico, Cuba, and Spain. Children from those different ethnic and cultural backgrounds may be more dissimilar among themselves than they are from children of other ethnic categories such as African-Americans or Caucasians.

Under the selection model used by this school district, each of the children to be certified gifted must have achieved a score on a nationally standardized group achievement test in the 90th percentile or higher. Since not every child is referred for evaluation, several sources of referral bias may begin at this stage of the process (e.g., based on gender, culture, or verbal skill level). Each then was further evaluated with a nationally standardized individual test of intelligence. Children were subsequently certified gifted in one of two ways: 1) an IQ score two standard deviations above the national mean or higher (e.g., WISC-R FSIQ $\geq$ 130), or 2) an individual IQ score $\geq$ 120 plus the presence of two or more identified areas of risk, as discussed above. An examination of the risk factors demonstrated considerable heterogeneity within each ethnic group and across ethnic groups, as would be expected (see Figure 1). Again, problems in the use of self-report data can be seen. Certain groups may tend to under-report, and teachers may tend to selectively report factors seen more frequently in one group than in another (e.g., language, which is especially obvious without much depth of knowledge about the child or family).

Figure 1. Within each ethnic group, the percentage at each level of risk in the population of children referred as potentially gifted.



Each child in this study was given the Wechsler Intelligence Scale for Children-Revised (WISC-R) by a school psychologist as part of the evaluation process between 1984 and 1991. The two supplemental subtests (Mazes and Coding) were not routinely administered in this district, and the Comprehension and Digit Span subtests were given to too few of the children to be included in multivariate analyses. The omission of Comprehension and Digit Span for so many children introduced another possible source of bias, in that prorated IQ scores were used for those children and may not represent the same Full Scale score as would have resulted from the inclusion of all subtests. Furthermore, the decision to administer those two subtests to some but not all children may have been based on systematic differences in attributes such as verbal facility and/or cultural and language differences.

Study I. For the study of observed base rates of VIQ-PIQ differences, the sample included every African-American, Asian, Caucasian, Filipino, and Hispanic child who achieved a Full Scale IQ score of at least 130 (by definition, two standard deviations above the mean) on the WISC-R between the years 1984 and 1991, inclusive. Forty six percent were female. Ethnic composition of the sample is summarized in Table 1.

Table 1. *Composition of the VIQ-PIQ base rate sample*

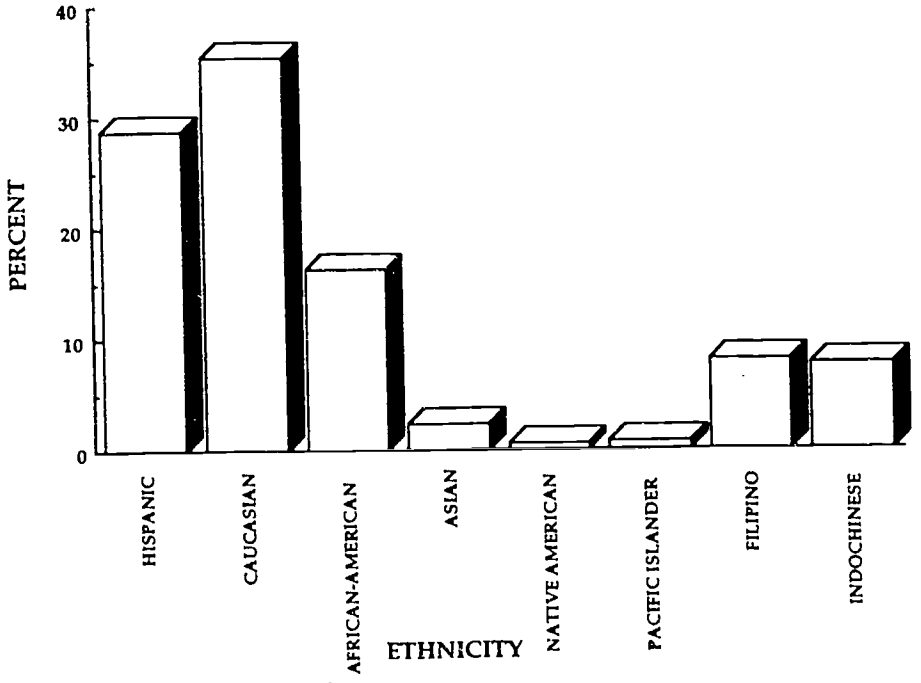
		Percent of each group at each Level of Risk		
Group	Number	0	1	>1
African-American	252	52	20	28
Asian	202	53	20	27
Caucasian	4895	71	16	12
Filipino	182	40	20	40
Hispanic	265	49	14	37
Total	5796			

Heterogeneity of levels of risk for these children, within and across ethnic groups, can be seen in Figure 2. Comparison of Figures 1 and 2 strikingly demonstrates the WISC-R disadvantage associated with a high risk environment, for children of every ethnic background. In each ethnic group, children from high risk environments differentially tended to score below 130 in FSIQ on the WISC-R and so were selected out of the sample for the base rate study. Use of a single cut-off score by a school district would obviously tend to exclude those children from enrichment programs as well.

Figure 2. Within each ethnic group, the percent at each level of risk in the population of children with FSIQ at least 130.

Study II. 19,826 children were identified as potentially gifted by the San Diego City School District in the years from 1984 through 1991. A total of 8396 children were subsequently administered the WISC-R, while others were evaluated with other instruments such as the Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983). From the group administered the WISC-R, a random sample of 1438 (713 female) was chosen to be ethnically proportionate to the composition of the district population in the academic year 1990-1991 (see Figure 3). The random sample consisted of 258 African-American, 36 Asian, 560 Caucasian, 128 Filipino, and 456 Hispanic children. Size of the sample was limited by the proportionately small number of Hispanic children administered the WISC-R, as compared to their numbers in the district population. In its determination to find equitable selection methods, this school district uses tests other than the WISC-R whenever possible with the predominantly Spanish-speaking members of its large population of Hispanic children.

Figure 3. San Diego Unified School District ethnic composition, 1991/1992.



III. Results

Study I: Base Rates for VIQ-PIQ Differences

Descriptive statistics for the WISC-R scores of the sample of 5796 children with Full Scale IQ values of at least 130 are presented in Table 2.

Table 2. *Verbal, Performance, and Full Scale Scores as a Function of Ethnic Group*

Group	VIQ Mean	(*sd*)	PIQ Mean	(*sd*)	FSIQ Mean	(*sd*)
African-American	136 .2	(8.02)	129.6	(9.00)	136.7	(6.09)
Asian	135.0	(9.97)	136.6	(8.78)	139.8	(7.18)
Caucasian	136.4	(8.55)	132.3	(9.08)	138.4	(6.56)
Filipino	132.3	(9.63)	134.5	(8.92)	137.3	(6.18)
Hispanic	135.2	(8.89)	133.6	(7.95)	138.4	(6.33)

Preliminary analyses were conducted to examine the trends in this group of intellectually gifted children. Gender effects were analyzed in a 2 (Gender) by 3 (Test Score) mixed repeated measures analysis of variance. Significant main effects were found for Gender, $F(1, 5794) = 53.67$, $p < .001$, but there was no interaction effect. Boys, on the average, scored higher than girls, as can be seen in Table 3. Given the standard error of measurement of the WISC-R, although the differences were statistically significant, they were clinically irrelevant.

Table 3. *Verbal, Performance, and Full Scale Scores as a Function of Gender*

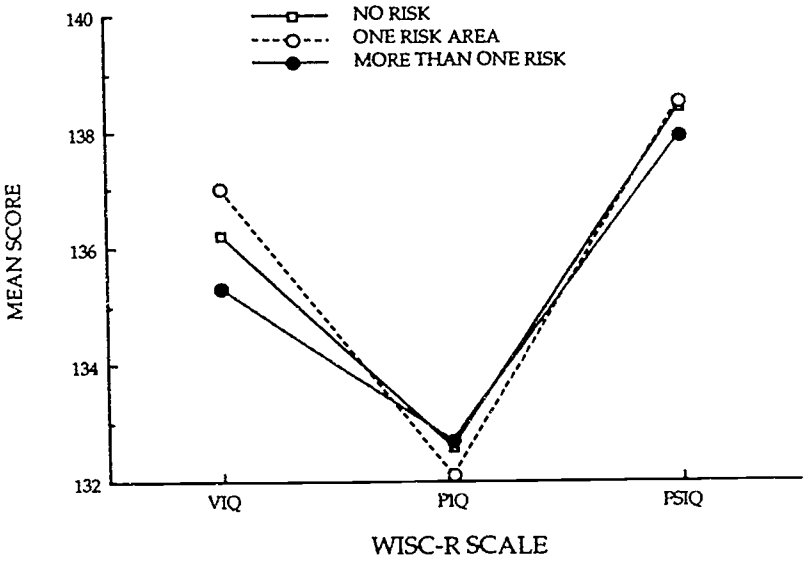
Gender	VIQ Mean	(*sd*)	PIQ Mean	(*sd*)	FSIQ Mean	(*sd*)
Female	135.5	(8.39)	131.9	(8.83)	137.6	(6.24)
Male	136.8	(8.85)	133.0	(9.25)	138.9	(6.77)

Verbal, Performance, and Full Scale IQ values for each ethnic group were analyzed in a 5 (Ethnicity) X 3 (Test Score) mixed repeated measures analysis of variance. Results revealed significant main effects for Ethnicity, $F(4, 5791) = 7.41$, $p < .001$ and for Test Score, $F(2, 5791) = 200.54$, $p < .001$. These main effects must, however, be interpreted in light of the significant Ethnicity by Test Score interaction, $F(8, 5791) = 25.70$, $p < .001$. As can be seen in Table 2 and confirmed in post hoc multiple Scheffé comparisons, Filipino children were significantly lower in Verbal IQ scores than African-American, Caucasian, or Hispanic children. On the other hand, Filipinos were higher in Performance IQ than African-Americans or Caucasians, and African-Americans were lower than any other group. Clear differences in pattern of strengths and weaknesses among these gifted children seem apparent.

To investigate the possibility that observed ethnic group differences in Verbal IQ scores could be due primarily to differences in risk status, a oneway analysis of covariance was performed with level of risk as the covariate. Results indicated that level of risk was a nonsignificant covariate, and that ethnic status remained a significant effect, $F(4,5790) = 11.50$, $p < .001$, regardless of risk.

In a series of oneway analyses of variance, level of identified risk was found to have a significant effect only on Verbal IQ scores, $F(2, 5793) = 9.46$, $p < .001$, but not on Performance or Full Scale IQ scores. Post hoc Scheffé comparisons revealed that those with one and only one identified area of risk obtained Verbal scores significantly higher than those with no risk, whose scores were higher than those with more than one risk area (see Figure 4).

Figure 4. WISC-R scores as a function of level of risk.



Thus the presence of multiple areas of risk or hardship in a gifted child's environment appears to be associated with lower performance on the Verbal Scale of the Wechsler, while the presence of one unspecified risk factor alone does not.

In an effort to understand the finding that children with one and only one risk had higher mean VIQ than those at no risk, a series of hypotheses was tested. The first hypothesis was that, among children with a single identified risk, either ethnic groups with higher mean VIQ (i.e., Asians and Caucasians) or males (who had higher VIQ than females) were disproportionately highly represented. A oneway analysis of variance compared VIQ in the two levels of risk, with ethnic group membership and gender as covariates. Ethnicity was a significant covariate, $F(1,4892) = 9.39$, $p < .01$, as was gender, $F(1, 4892) = 27.45$, $p < .001$. Risk level, however, remained a significant main effect, $F(1,4892) = 6.87$, $p < .01$. Therefore, the VIQ differences across level of risk did not appear to be a simple function of ethnicity or gender alone. In fact, a 2 (Risk Level) X 2 (Gender) X 5 (Ethnicity) ANOVA demonstrated significant main effects for Level of Risk, $F(1,4876) = 7.27$, $p < .01$, Gender, $F(1,4876) = 27.78$, $p < .001$, and Ethnicity, $F(1,4876) = 6.43$, $p < .001$. None of the interaction effects were significant. Since neither gender nor ethnic background accounted for the differences in VIQ across risk, an alternative hypothesis that type of risk accounted for the higher mean in one-risk children was investigated in a oneway analysis of covariance with type of risk as the covariate. Type of risk was a significant covariate, $F(1,4893) = 9.24$, $p < .01$. When the variance accounted for by type of risk was removed, level of risk was no longer a significant effect. To further elucidate this finding, frequencies of ethnic background and gender across type of risk were examined in children with only one area of risk. Most frequent was

emotional risk (30.5% of the total), followed by health (26.1% of the total) and environmental (19.0% of the total). The presence of cultural, economic, or language hardship alone was relatively rare (2.2%, 10.1%, and 11.7%, respectively). Within the group at emotional risk, 89.8% of the children were Asian and Caucasian; 53.0% were male. Among those at health risk, 96.0% were Asian and Caucasian; 64.9% were male. In the environmental risk group, 89.1% were Asian and Caucasian; 56.8% were male. Type of risk appears to be a mediator for ethnicity and for gender, and the higher mean VIQ scores in children with only one risk area appear to be explainable in terms of a higher proportion of males and of Asians and Caucasians (all associated with higher mean VIQ) in the group of children identified with only emotional, health, or environmental risk than in the overall sample.

Base rates for the difference between Verbal and Performance IQ score were obtained and are summarized in Tables 4 through 8. Ranges were defined to be consistent with those of Matarazzo and Herman (1984), so that comparisons with their findings could be made.

Table 4. *African-Americans: Cumulative Percentage Distributions of the Difference Between WISC-R VIQ and PIQ*

Size of the Difference Between VIQ and PIQ	%V>P (+ Difference)	%P>V (-Difference)	Sum of WISC-R + and - Differences	Cumulative Percentage
30 and above	3.97	0	3.97	100.00
26-29	3.57	0	3.57	96.03
22-25	5.56	1.98	7.54	92.46
19-21	5.16	.79	5.95	84.92
16-18	8.33	1.19	9.52	78.97
13-15	6.35	1.19	7.54	69.45
10-12	5.56	3.17	8.73	61.91
7-9	9.92	4.37	14.29	53.18
4-6	6.75	7.94	14.69	38.89
1-3	11.11	9.92	21.03	24.20
0	—	—	—	3.17

Table 5. *Asians: Cumulative Percentage Distributions of the Difference Between WISC-R VIQ and PIQ*

Size of the Difference Between VIQ and PIQ	%V>P (+ Difference)	%P>V (-Difference)	Sum of WISC-R + and - Differences	Cumulative Percentage
30 and above	0	.99	.99	100.00
26-29	.99	1.49	2.48	99.05
22-25	2.48	3.96	6.44	96.57
19-21	2.48	2.97	5.45	90.13
16-18	3.47	4.95	8.42	84.68
13-15	4.46	4.46	8.92	76.26
10-12	5.94	7.43	13.37	67.34
7-9	5.94	10.40	16.34	53.97
4-6	9.9	8.91	18.31	37.63
1-3	8.91	6.44	15.35	18.82
0	—	—	—	3.47

Table 6. *Caucasians: Cumulative Percentage Distributions of the Difference Between WISC-R VIQ and PIQ*

Size of the Difference Between VIQ and PIQ	%V>P (+ Difference)	%P>V (-Difference)	Sum of WISC-R + and - Differences	Cumulative Percentage
30 and above	2.35	.16	2.51	100.00
26-29	2.25	.57	2.82	97.51
22-25	3.49	1.41	4.90	94.69
19-21	5.05	1.53	6.58	89.79
16-18	5.99	2.49	8.48	83.21
13-15	6.07	2.80	8.87	74.73
10-12	8.09	5.03	13.12	65.86
7-9	8.83	5.78	14.61	52.74
4-6	8.95	7.84	16.79	38.13
1-3	9.85	8.34	18.19	21.34
0	—	—	—	3.15

Table 7. *Filipinos: Cumulative Percentage Distributions of the Difference Between WISC-R VIQ and PIQ*

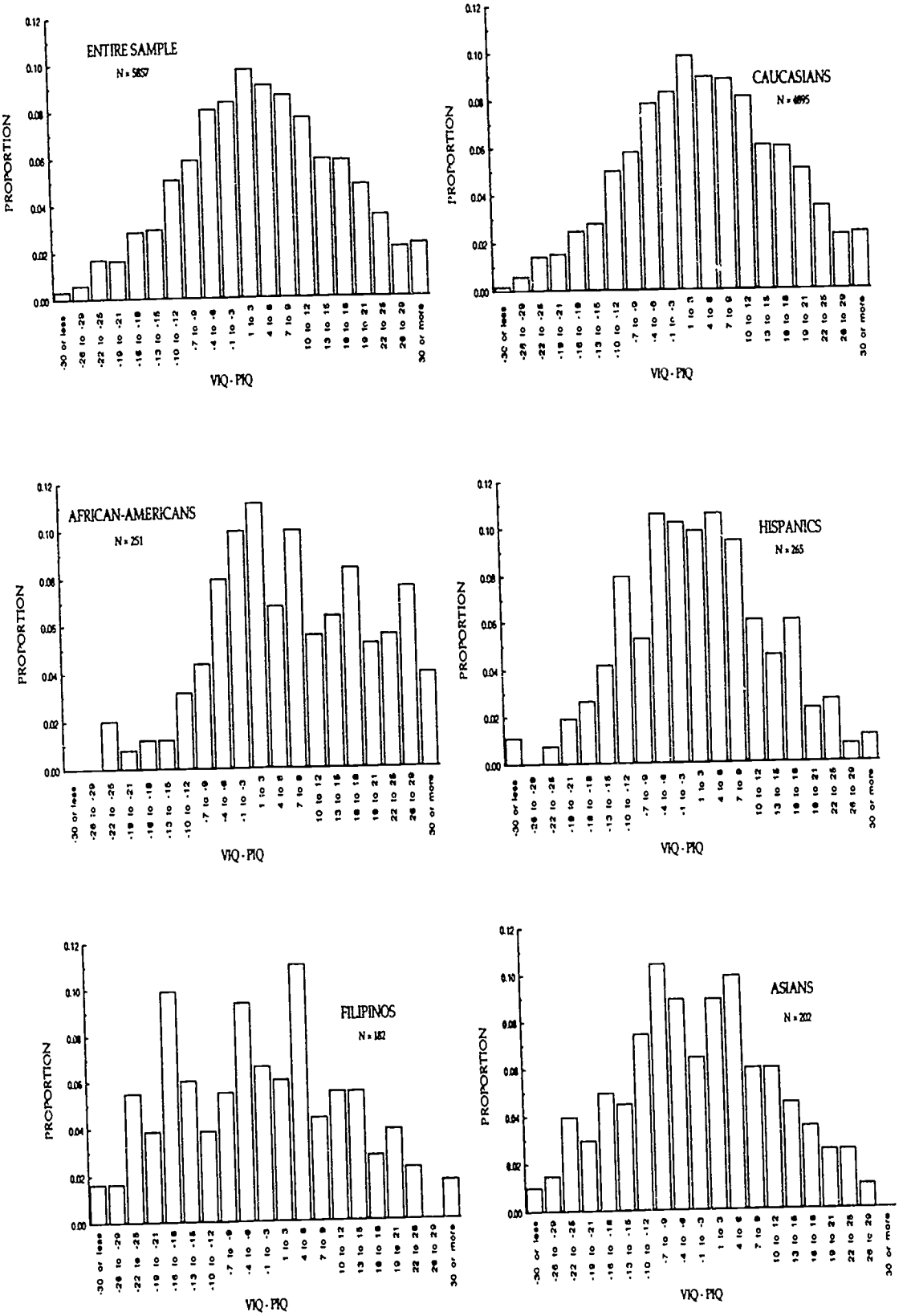
Size of the Difference Between VIQ and PIQ	%V>P (+ Difference)	%P>V (-Difference)	Sum of WISC-R + and - Differences	Cumulative Percentage
30 and above	1.65	1.65	3.30	100.00
26-29	0	1.65	1.65	96.70
22-25	2.20	5.49	7.69	95.06
19-21	3.85	3.85	7.70	87.37
16-18	2.75	9.89	12.64	79.67
13-15	5.49	6.04	11.53	67.03
10-12	5.49	3.85	9.34	55.50
7-9	4.40	5.49	9.89	46.16
4-6	11.00	9.34	20.34	36.27
1-3	6.04	6.59	12.63	15.93
0	—	—	—	3.30

Table 8. *Latinos/Hispanics: Cumulative Percentage Distributions of the Difference Between WISC-R VIQ and PIQ*

Size of the Difference Between VIQ and PIQ	%V>P (+ Difference)	%P>V (-Difference)	Sum of WISC-R + and - Differences	Cumulative Percentage
30 and above	1.13	1.13	2.26	100.00
26-29	.75	0	0.75	97.72
22-25	2.64	.75	3.39	96.97
19-21	2.26	1.89	4.15	93.58
16-18	6.04	2.64	8.68	89.43
13-15	4.53	4.15	8.68	80.75
10-12	6.04	7.92	13.96	72.07
7-9	9.43	5.28	14.71	58.11
4-6	10.57	10.57	21.14	43.40
1-3	9.81	10.19	20.00	22.26
0	—	—	—	2.26

Inspection of these tables suggests striking differences between ethnic groups. To examine those differences, the VIQ - PIQ frequency distribution for each ethnic group was compared to a reference distribution using a Chi Square test with 20 degrees of freedom. The reference distribution chosen was that of the standardization sample for the Wechsler Adult Intelligence Scale-Revised, reported by Matarazzo and Herman (1984), since those authors reported direction as well as magnitude of the VIQ-PIQ difference. Hispanics and Asians were not found to differ from the WAIS-R standardization sample. African-Americans, $\chi^2(20, N = 252) = 196.9$, $p<.001$, Caucasians, $\chi^2(20, N = 4895) = 1382.6$, $p<.001$, and Filipinos, $\chi^2(20, N = 182) = 90.7$, $p<.001$, did differ significantly from the reference distribution. The nature of those distributions is shown in Figure 5.

Figure 5. Distributions of (VIQ - PIQ) differences as a function of ethnic group.



Again we see the trend for African-American and Caucasian children to have higher VIQ than PIQ, while the reverse is true for Filipino children. The previous analyses found the trend in group mean scores. The Chi Square differences between VIQ and PIQ confirm those findings in individuals and further strengthens the evidence for differences in patterns of strengths and weaknesses across ethnic background.

Given these differences in distributions between ethnic groups, it becomes crucial to look at population incidence of large VIQ-PIQ discrepancies as a function of ethnic background. Only in this way can we determine whether an event that is rare in one group, and is taken as a clinical indicator of abnormality, also holds for other groups. Within each ethnic group, occurrences of magnitudes of VIQ-PIQ discrepancies were counted so that population rarity could be compared with statistical significance (as presented in the WISC-R manual) for each group. That is, a VIQ-PIQ difference of 12 points has been found to be statistically significant at the .05 level. This finding is frequently misinterpreted to mean that only about 5% of normal children will have a difference of that magnitude. However, Kaufman (1976) pointed out that approximately 30% of normal children with average intelligence have discrepancies at least that high, as do 36% of children in the standardization sample with IQ scores of at least 120. In Table 9, the difference required for statistical significance is compared to that actually observed in each of the ethnic groups. For example, a difference of 12 points is needed to be sure (within an error probability of .05) that a child's Verbal and Performance abilities are significantly different. For the Asian children in this sample, however, a difference of 25 points or more is needed in order for the difference to be rare enough to be observed only about 5 percent of the time.

Table 9. *Empirically Different Magnitudes of VIQ-PIQ Discrepancies, as a Function of Ethnic Group*

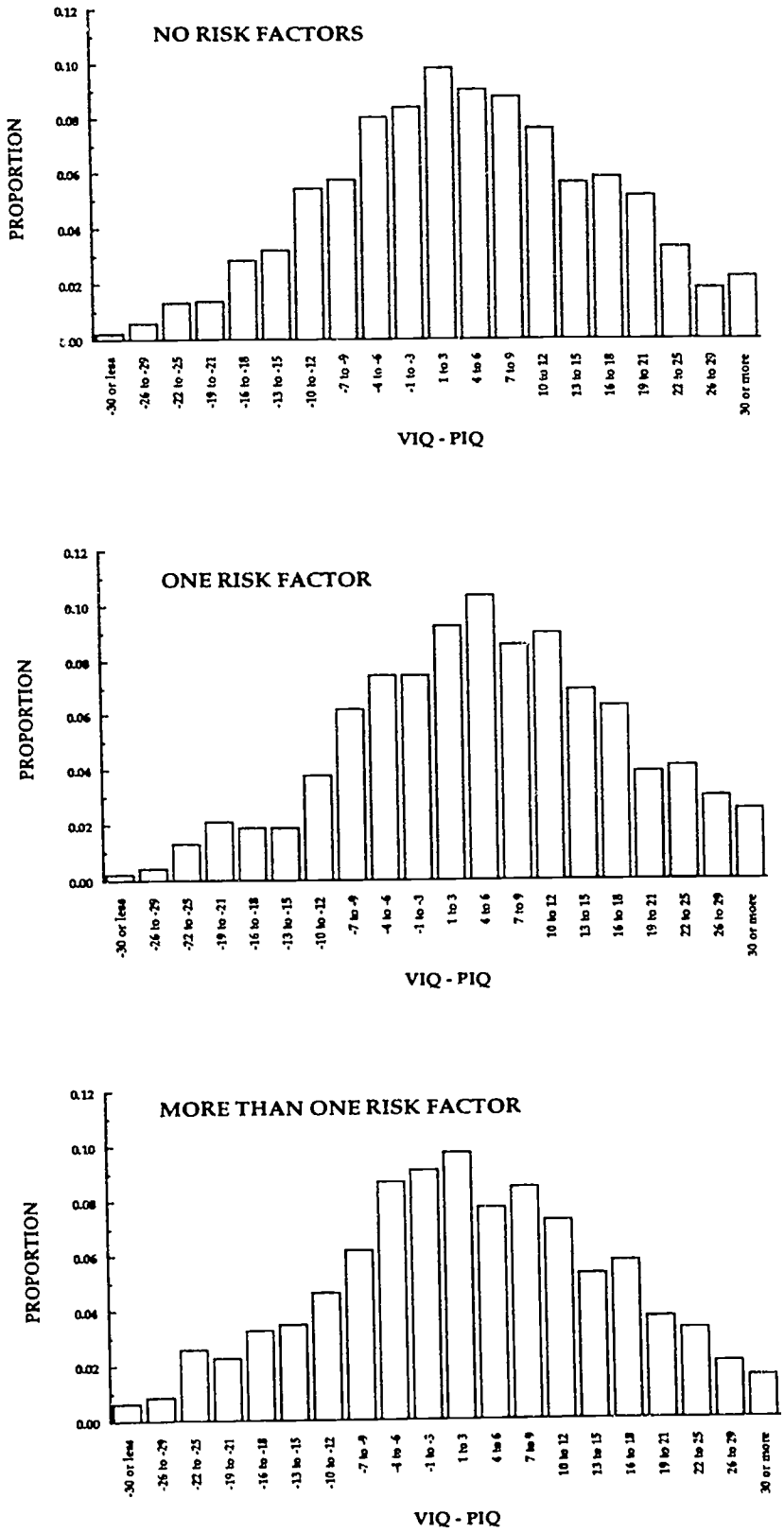
	Magnitude of Difference Required Statistically*	Magnitude of Difference Empirically Observed at Each Level of Probability				
p value		Caucasian	Hispanic	Afr.Amer.	Asian	Filipino
.15	8	20	18	22	19	21
.10	10	23	19	25	22	23
.05	12	27	23	29	25	27
.01	15	34	40	35	30	38

* to be reliably different from 0

Considerable variation between groups can be seen in Table 9. Although a VIQ-PIQ difference of 30 points is rare in the gifted Asian population of our sample (occurring only about once in every one hundred children), 5 in every one hundred African-American children are observed to have that difference, and even more children in each of the other three groups. One can easily imagine a scenario in which, for example, norms are set using a predominantly Asian population, rare (less than 5% of the population) VIQ-PIQ differences are defined to be a diagnostic indicator for learning disabilities, and that standard is used for all children. In this particular gifted sample, such a criterion could lead to labelling twice as many Caucasian, Filipino, and African-American children as Asian or Hispanic learning disabled. The scenario is admittedly an exaggerated one and it is to be hoped that in actual practice one single test is never the sole criterion for diagnostic or placement decisions.

The effect of level of risk was further examined as base rates among children at no, low, and high (2 or more identified areas of risk) risk were compared (see Figure 6).

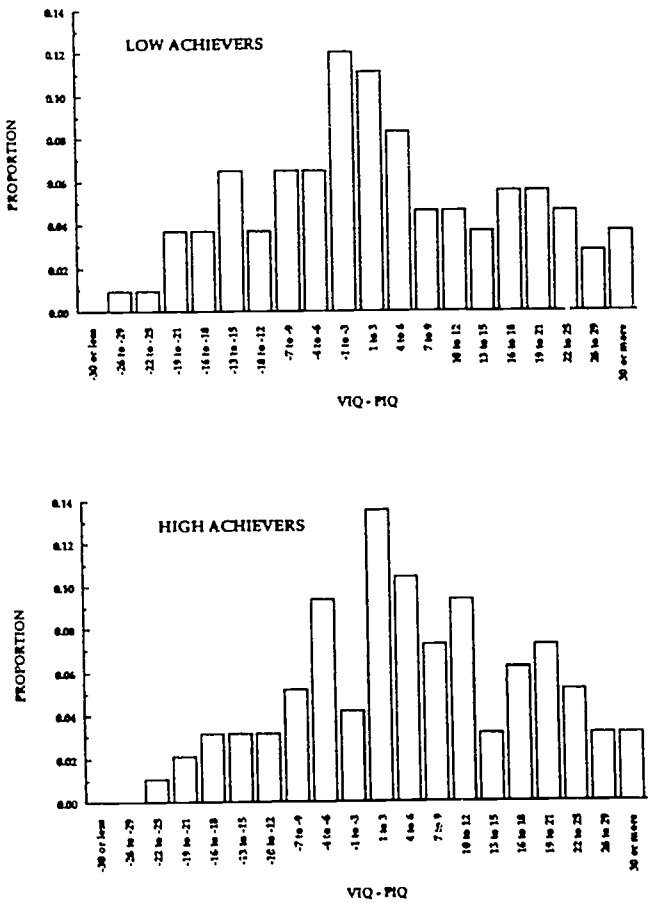
Figure 6 Distributions of (VIQ - PIQ) differences as a function of level of risk.



Distributions were statistically compared using Kolmogorov-Smirnov 2-sample tests. Children with no identified risk differed from those with one risk ($Z = 1.526$, $p < .05$), and children with only one risk differed from those with more than one risk ($Z = 1.951$, $p < .001$). As can be seen in the figure, children from high risk backgrounds more frequently tended to have higher PIQ than VIQ. This comes as no surprise, in light of findings from group means that children at high risk have lower mean VIQ. Each of these children achieved a Full Scale IQ of at least 130. In order to accomplish that in the face of a disadvantaged VIQ, PIQ must be even higher than for those at no risk. Again, we see differences in pattern across groups.

Finally, subsamples of this demonstratedly gifted sample were selected so that rates of VIQ-PIQ differences could be compared in gifted high and low achievers. For this purpose, scores on the California Test of Basic Skills (CTBS) were obtained. Two subsamples were selected; 96 children whose CTBS scores were all at a stanine of 9 were designated "high achievers", and 108 children whose CTBS scores were all at a stanine of 6 or below were called "low achievers". Single classification ANOVAs revealed that the groups did not significantly differ in PIQ; for high achievers, $M = 132.2$ (SD=9.7), while for low achievers M=130.6 (SD=9.1). Low achievers were, however, significantly different from high achievers in VIQ, $F(1,203) = 13.49$; $p < .001$. Group means were 137.8 (SD=8.5) and 133.4 (SD=7.8), respectively.

Figure 7. Distribution of VIQ-PIQ differences at the extremes of achievement.



VIQ - PIQ distributions for the two groups are shown in Figure 7. No significant differences were found between the two distributions (Kolmogorov-Smirnov $Z = 1.007$, $p = .263$). This implies that use of large VIQ-PIQ discrepancy as an indicator of risk for low achievement is indeed fallacious, since relatively large VIQ-PIQ discrepancies are as likely to be seen in high achievers as in low achievers.

Study II: Multivariate attempts to use the WISC-R to select an ethnically balanced gifted population.

An ethnically balanced random sample of 1438 children identified as potentially gifted was generated. Mean scores on the eight subtests of the WISC-R routinely administered in the district are summarized in Table 10.

Table 10
WISC-R Scores in a Randomly Balanced Sample of Children Identified Potentially Gifted

Score	Entire Sample	African-American	Asian	Caucasian	Filipino	Hispanic
FSIQ	130.6*	127.0	134.8	133.1	128.5	129.8
	(11.3)**	(11.9)	(10.2)	(10.0)	(11.4)	(11.8)
VIQ	129.1	127.5	129.8	132.0	124.1	127.8
	(12.1)	(12.1)	(12.4)	(11.0)	(12.5)	(12.5)
PIQ	125.6	120.4	132.8	127.1	127.2	125.6
	(12.3)	(12.8)	(11.6)	(11.4)	(12.3)	(12.3)
Information	13.5	12.9	14.1	14.1	12.9	13.3
	(2.5)	(2.4)	(3.1)	(2.4)	(2.4)	(2.6)
Similarities	15.8	15.7	15.2	16.2	14.8	15.8
	(2.4)	(2.5)	(2.5)	(2.2)	(2.7)	(2.4)
Arithmetic	13.9	13.5	14.4	14.3	13.6	13.8
	(2.4)	(2.4)	(2.3)	(2.3)	(2.3)	(2.4)
Vocabulary	14.7	14.6	14.9	15.2	13.8	14.2
	(2.6)	(2.5)	(2.6)	(2.4)	(2.8)	(2.7)
Picture Completion	13.2	12.7	13.5	13.3	13.0	13.4
	(2.4)	(2.3)	(2.2)	(2.4)	(2.4)	(2.4)
Picture Arrangement	14.0	13.3	14.5	14.2	14.0	14.1
	(2.7)	(2.7)	(2.7)	(2.7)	(3.0)	(2.7)
Block Design	14.0	12.7	16.2	14.4	14.7	13.8
	(3.0)	(3.0)	(2.8)	(2.8)	(2.7)	(2.9)
Object Assembly	13.3	12.4	13.8	13.6	13.3	13.3
	(2.8)	(2.9)	(2.8)	(2.6)	(2.9)	(2.7)

* Mean
** Standard Deviation

Inspection of Table 10 reveals the problem experienced by any diverse school district in its efforts to provide equal access to gifted programs based primarily on Full Scale IQ as measured by the WISC-R. As has happened in San Diego City Schools, Caucasians and Asians will be over-represented, while Hispanics and African-Americans will be under-represented. Assuming that the WISC-R does indeed predict academic achievement and that an ethnic balance in gifted programs is a desirable and in fact necessary goal, each ethnic subsample was divided on the basis of FSIQ: the upper 70% of each group was designated "gifted" for the purposes of the following analyses, and the lower 30% of each was designated "nongifted". Those percentages were estimated based on the overall number of children referred for individual testing versus the 70% finally selected for inclusion in gifted enrichment classrooms.

Stepwise multiple linear regression analyses were carried out on the scaled scores of the whole sample, as well as each ethnic subsample, in order to determine which subtests of the WISC-R best predict giftedness for each group. Results are summarized in Table 11.

Table 11. *Stepwise Multiple Linear Regression Models to Predict Giftedness*

Sample	Subtests in the Model	ß	*R*	R^2
Entire Sample	Vocabulary	.156		
	Object Assembly	.172		
	Picture Completion	.153		
	Similarities	.167		
	Block Design	.150		
	Picture Arrangement	.144		
	Arithmetic	.105		
	Information	.101	.696	.485
African-Americans	Information	.180		
	Object Assembly	.189		
	Similarities	.163		
	Block Design	.150		
	Picture Arrangement	.140		
	Arithmetic	.131		
	Vocabulary	.114	.659	.483
Asians	Information	.457		
	Block Design	.357	.604	.365
Caucasians	Object Assembly	.236		
	Vocabulary	.181		
	Picture Completion	.183		
	Arithmetic	.159		
	Block Design	.172		
	Similarities	.157		
	Picture Arrangement	.120		
	Information	.112	.753	.567
Filipinos	Picture Arrangement	.291		
	Picture Completion	.271		
	Block Design	.284		
	Similarities	.274	.699	.489
Hispanics	Similarities	.226		
	Object Assembly	.168		
	Information	.154		
	Picture Completion	.151		
	Picture Arrangement	.154		
	Vocabulary	.161		
	Block Design	.131	.735	.533

For each group except Asians, the best stepwise selection model was able to account for approximately 50% of the variance or more: for Asians, R^2 was only .36. Variables in the model differed across ethnic groups, as well. For Caucasians, as for the sample as a whole, all subtests entered into the equation. For Hispanics, only Arithmetic failed to enter, while for African-Americans the Picture Completion subtest did not enter the model. The best-fitting model for Filipinos included Picture Arrangement, Picture Completion, Block Design, and Similarities. Only two subtests were include in the model for Asians: Information and Block Design. Again we see differences in pattern of strengths and weaknesses, reflected in different predictors of giftedness across ethnic background.

To determine the efficacy of the best predictive model, discriminant analysis was performed for the entire sample using all subtests as predictors and giftedness as the criterion. With two groups (gifted and non-gifted), one discriminant factor was generated. Results are summarized in Table 12.

Table 12. *Discriminant Function Coefficients for the Identification of Giftedness in the Entire Ethnically Balanced Sample*

Standardized Coefficients		Pooled Correlations Between Subtests and the Discriminant Function	
Information	.252	Similarities	.566
Similarities	.376	Vocabulary	.555
Arithmetic	.111	Information	.549
Vocabulary	.265	Object Assembly	.461
Picture Completion	.270	Block Design	.437
Picture Arrangement	.273	Picture Completion	.424
Block Design	.225	Picture Arrangement	.412
Object Assembly	.298	Arithmetic	.347

False positives, false negatives, and hit rates, in percentages, are provided for the whole sample and for each ethnic group within that sample in Table 13.

Table 13. *Hit Rates, False Positives, and False Negatives for the Best Overall Model*

Group	Hit Rate	False Positives*	False Negatives**
Entire Sample	89.3	7.5	19.9
African-Americans	81.7	0	23.6
Asians	93.9	33.3	0
Caucasians	87.9	42.9	.5
Filipinos	89.6	3.6	12.4
Hispanics	95.0	4.2	5.3

* of those who were not gifted, the percent called "gifted"
** of those who were gifted, the percent called "nongifted" by the model

The most critical errors are represented in the "false negatives" column of the table. Those numbers represent children who have unusually high potential that would not be recognized. Those children would be denied a chance to excel in special programs for the intellectually gifted. When we examine false positive and negative rates for subgroups, we see the repetitive pattern of over-selection of Caucasians and Asians accompanied by the under-selection of African-Americans. The one group for whom this model is an improvement is Hispanics. **Thus we demonstrate that no one single model using the WISC-R, no matter how sophisticated and complex, will select an ethnically proportionate sample for inclusion into enrichment programs for the gifted.**

The efficacy of individual models of selection, based on ethnic background, was investigated by performing discriminant analyses on each ethnic subsample, using all available subtests of the WISC-R as predictors and giftedness as the criterion. Results are summarized in Tables 14 through 18.

Table 14. *African-Americans: Discriminant Function Coefficients for the Identification of Giftedness*

Standardized Coefficients		Pooled Correlations Between Subtests and the Discriminant Function	
Information	.298	Information	.573
Similarities	.287	Object Assembly	.549
Arithmetic	.216	Arithmetic	.522
Vocabulary	.195	Similarities	.517
Picture Completion	.132	Block Design	.512
Picture Arrangement	.254	Vocabulary	.500
Block Design	.267	Picture Arrangement	.455
Object Assembly	.310	Picture Completion	.364
Eigenvalue .9512		Wilks' Lambda .513	

Table 15. *Asians: Discriminant Function Coefficients for the Identification of Giftedness*

Standardized Coefficients		Pooled Correlations Between Subtests and the Discriminant Function	
Information	.513	Information	.648
Similarities	.011	Vocabulary	.579
Arithmetic	.174	Arithmetic	.554
Vocabulary	.247	Picture Completion	.502
Picture Completion	.213	Block Design	.500
Picture Arrangement	.268	Similarities	.379
Block Design	.556	Object Assembly	.311
Object Assembly	-.055	Picture Arrangement	.209
Eigenvalue .746		Wilks' Lambda .573	

Table 16. *Caucasians: Discriminant Function Coefficients for the Identification of Giftedness*

Standardized Coefficients		Pooled Correlations Between Subtests and the Discriminant Function	
Information	.202	Object Assembly	.480
Similarities	.289	Vocabulary	.453
Arithmetic	.294	Information	.436
Vocabulary	.324	Block Design	.419
Picture Completion	.337	Picture Completion	.392
Picture Arrangement	.227	Similarities	.392
Block Design	.313	Arithmetic	.386
Object Assembly	.417	Picture Arrangement	.331
Eigenvalue 1.310		Wilks' Lambda .433	

Table 17. *Filipinos: Discriminant Function Coefficients for the Identification of Giftedness*

Standardized Coefficients		Pooled Correlations Between Subtests and the Discriminant Function	
Information	.003	Picture Arrangement	.507
Similarities	.321	Vocabulary	.486
Arithmetic	.172	Similarities	.471
Vocabulary	.199	Block Design	.468
Picture Completion	.419	Object Assembly	.427
Picture Arrangement	.435	Picture Completion	.424
Block Design	.413	Information	.376
Object Assembly	.230	Arithmetic	.359
Eigenvalue 1.076		Wilks' Lambda .482	

Table 18. *Hispanics: Discriminant Function Coefficients for the Identification of Giftedness*

Standardized Coefficients		Pooled Correlations Between Subtests and the Discriminant Function	
Information	.252	Similarities	.566
Similarities	.376	Vocabulary	.555
Arithmetic	.111	Information	.549
Vocabulary	.265	Object Assembly	.461
Picture Completion	.270	Block Design	.437
Picture Arrangement	.273	Picture Completion	.424
Block Design	.225	Picture Arrangement	.412
Object Assembly	.298	Arithmetic	.347
Eigenvalue 1.191		Wilks' Lambda .456	

Hit rates, false positive and false negative rates for the use of these individual functions are summarized in Table 19.

Table 19. *Hit Rates, False Positives, and False Negatives for the Best Individual Discriminant Function Models*

Model	Hit Rate	False Positives	False Negatives
African-American	93.6	17.5	3.1
Asian	90.9	16.7	7.4
Caucasian	95.0	12.2	2.3
Filipino	93.6	7.1	6.2
Hispanic	93.8	7.6	5.6

* of those who were not gifted, the percent called "gifted"
** of those who were gifted, the percent called "nongifted" by the model

By using individual models and capitalizing on differences in pattern of strengths and weaknesses across ethnic groups, rates have been improved for some groups. Caucasians and Asians are still overrepresented, as now are African-Americans. Hispanics and Filipinos have nearly equal false positive and false negative rates. An important note is that the models used to obtain these rates are based on functions that are *weighted* sums of subtests, and not simple combinations of subtests providing easy cut-off scores. These are the best rates available, based on fairly complex linear combinations. Any combination of subtest cutoff scores used in actual practice would necessarily have lower success rates.

To investigate the possibility that discrimination of nongifted from gifted could be improved by grouping those with similar patterns of abilities, African-Americans and Caucasians were considered together in one discriminant model. Both groups had a tendency for higher VIQ than PIQ (see baserate study, above). However, Caucasians are traditionally overselected and African-Americans underselected. Results of the analysis are presented in Table 20.

Table 20. *Discriminant Function Coefficients for Groups whose VIQ Exceeds PIQ (African-Americans and Caucasians)*

Standardized Coefficients		Pooled Correlations Between Subtests and the Discriminant Function	
Information	.150	Object Assembly	.542
Similarities	.290	Vocabulary	.526
Arithmetic	.257	Information	.509
Vocabulary	.304	Similarities	.481
Picture Completion	.273	Block Design	.473
Picture Arrangement	.219	Arithmetic	.473
Block Design	.224	Picture Completion	.433
Object Assembly	.349	Picture Arrangement	.407

Eigenvalue .8867 Wilks' Lambda .530

Using this discriminant model, overall hit rates have gone down to 80.6% for African-Americans and 90.1 for Caucasians. The misses, as expected, favor Caucasians (34.7% false positives, .5% false negatives) and again disadvantage African-Americans (0 false positives, 19.4% false negatives). No manipulation will improve the rates from the best individual ethnic group discriminant models, obtained using all available subtests.

One might be tempted to argue that identification could be improved and gifted programs could be ethnically balanced more economically by choosing each group's strongest subtest and basing the decision on a cutoff score applied to a different subtest for each group. For example, as was seen in Table 10, African-Americans' highest mean scaled score was Information, Asians' was Block Design, and so on. Only 24.5% of African-Americans scored below 12 on Information and only 25% of Asians scored below 15 on Block Design. Therefore the same cutoff score would not work in both groups on the individually selected subtests.

In similar fashion, it might be proposed that there exists one subtest that, at a given cutoff score, would select a balanced population. Not only does that prove not to be the case, but a more fundamental issue is involved in this and in the proposal to use a different subtest for each group. The basic argument for using the WISC-R as a selection tool for intellectually gifted enrichment programs is that it in some way measures a broad array of abilities associated with achievement in school. By narrowing the test down, even to four subtests (much less one or two), the predictive power of the test is greatly diminished.

Thus we see that no single selection model using the WISC-R will result in an ethnically balanced sample of gifted children from this population. In fact, the most accurate and most complicated individual models for ethnic subgroups are not uniformly accurate either. There appears to be no way to use the WISC-R to derive cut-off inclusion/exclusion scores in this ethnically diverse sample for use in selecting balanced populations for gifted programs in the schools.

IV. Discussion

Intellectually gifted children show differences in the pattern of their strengths and weaknesses on the WISC-R, across ethnic background and across levels of risk in the environment. In the first phase of this work, an analysis of base rates of VIQ-PIQ differences in 5796 children with FSIQ scores at least two standard deviations above the mean (FSIQ $\geq$ 130) revealed that African-Americans and Caucasians tended to have a higher VIQ than PIQ, whereas in Filipino children the tendency was the reverse. These trends, evident in group data, were confirmed in frequency distributions of individual difference scores. The distributions of VIQ-PIQ difference scores of Asians and Hispanics most closely resembled those obtained by Matarrazzo (1984) from the standardization sample for the WAIS-R, and most closely approximated normal distributions.

Groups divided on the basis of level of risk from factors such as significant health problems, economic hardship, emotional deprivation, or cultural and linguistic factors were also found to differ in pattern of strengths and weaknesses. Such hardships proved to be consistently associated with lower VIQ relative to PIQ in the individual. It can be surmised that, of all children at risk from two or more of these factors, this sample contained only the most invulnerable children—only the children still able to achieve a Full Scale score two standard deviations above the mean—and that in a randomly selected population across IQ ranges, the differences would be more extreme. Comparison of the relatively low proportion of high-risk children seen in the gifted base rate sample, as opposed to the entire sample of children referred for giftedness assessment, appears to corroborate that hypothesis (refer to Figures 1 and 2).

The myth that relatively large VIQ-PIQ discrepancies are somehow a diagnostic indicator for learning difficulties, such as low achievement relative to potential, was debunked in this sample. Groups of gifted children at the extremes of achievement (all achievement scores in the ninth stanine versus all achievement scores in the sixth stanine or lower) were compared and found to have equivalent ranges of VIQ-PIQ difference scores.

In the second phase of the work, an ethnically proportionately balanced sample of 1438 potentially gifted children was randomly selected. From that sample, selection models were derived and examined for goodness of fit in an effort to find a way to use the WISC-R to select a balanced population for educational enrichment programs. Alternative hypotheses that 1) there exists a single pattern of subtest scores that predicts giftedness equally across ethnic background, or 2) there is a unique pattern of cognitive strengths and thus different predictors of giftedness across groups, were investigated. The best single model obtained by discriminant analysis appeared accurate overall. When examined in terms of individual ethnic groups, however, it proved to be biased in favor of Caucasians and Asians, and biased against African-Americans and Filipinos. As was seen in the base rate study, different patterns of strengths were evident across groups. Even accounting for those differences with individual best-fitting models, efforts to improve selection balance failed. The very best individual models overselected African-American, Asian, and Caucasian children. **No way was found to use the WISC-R to select a proportionately balanced population.** If individual subtests or combinations of two or more subtests are used, as is suggested by some authors (Dirks, Wessels, Quarforth, & Quenon, 1980; Elman, Blixt, & Sawicki, 1981; Karnes & Brown, 1981; Kaufman, 1979; Killian & Hughes, 1978; Sattler, 1988), discriminability suffers. Perhaps more importantly, predictive power of the WISC-R is decreased.

The present results confirm the findings of differences in pattern of WISC-R performances between ethnic groups reported by Saccuzzo et al. (1992). Moreover, for the first time, base rates for

VIQ-PIQ difference scores are presented for a large, ethnically diverse sample of gifted children. The distributions of VIQ-PIQ difference scores were found to be substantively different in shape as well as in direction across ethnic groups, in contrast to the report that discrepancy scores "did not vary too greatly with ... race" at any level of IQ in the standardization sample of the WISC-R (Kaufman, 1976). In fact, statistical power was too low for identification of differences had they existed in the higher end of IQ scores in the standardization sample. Also for the first time, high levels of risk in a child's social and home environment have been shown in the present work to be associated with lower VIQ relative to PIQ in children in the upper end of the IQ distribution. **Moreover, the influence of risk factors appears to confer a disadvantage over and above any effect of ethnic background.**

No model was found to enable the WISC-R to be used to select equal proportions of gifted children from a variety of ethnic backgrounds. Therefore, in an ethnically diverse population, it would seem that Sattler (1988) is correct in saying that children who are culturally different are difficult to identify. The results of commonly used identification practices can be seen nationwide in the over-representation of Asians and Caucasians, as well as the under-representation of African-Americans and Hispanics. **The results of this study strongly suggest that use of the WISC-R in diverse populations as the primary selection device for gifted programs is an inappropriate use of the test, if one of the goals of such use is to select proportionately representative numbers from each group.**

Uniformity assumption myths abound in psychological assessment. Sattler (1988) presented cogent arguments in favor of national norms and against the idea of pluralistic norms. He pointed out, in part, that the WISC-R was standardized on a carefully stratified sample with ethnic minorities represented in proportion to their representation in the population. That is certainly true on a national level, but the appropriateness of using national norms so derived to define cut-off scores in a population which is predominantly non-Caucasian, as is this school district, is questionable. Normative scores are derived based on factors known to affect test performance: the WISC-R manual provides only age-corrected norms. **The current study has demonstrated that differences in test performance on the WISC-R exist across ethnic background and across gender.** In the past, the major psychological and educational assessment devices have been standardized primarily in terms of age or education, and sometimes gender. More recently, authors have stressed the importance of differences and the need for sets of norms that consider multiple factors such as gender, age, and education concurrently (Heaton et al., 1986), especially when these scores are used for clinical decision-making. Use of the WISC-R for selection of individual children for enrichment programs is, in essence, a clinical decision-making process. The idea of pluralistic norms based on ethnic background, however, is politically an extremely sensitive issue. Perhaps the clearest conclusion from the findings of the present work is that the test, with existing norms, produces scores that are certainly more appropriate for some groups of children than for others.

If "intelligence" or "intellectual giftedness" were to be defined as exactly those abilities underlying the quality of an individual's performance on the WISC-R, then we would have to conclude that the WISC-R is the best instrument to use for selection, regardless of any socio-political considerations such as the need for ethnic, socioeconomic, or even gender balance. However, we are dealing with a theoretical construct (intelligence) imperfectly measured by the WISC-R within a known error of measurement. This work has demonstrated that groups divided either on ethnic background or on environmental factors differ in the pattern of their performances on the WISC-R. The results do not and can not address the issue of how much of the individual and group differences are a result of biologic (presumably neural) differences or of environmental influences. Aside from the social and political implications of the use of the WISC-R as an entry criterion in diverse populations, the results of the present work indicate that the basic assumption of uniformity of pattern of performance across groups on the WISC-R is flawed. For whatever reason, be it biologic, environmental, or a combination of the two, pattern of performance across groups is not uniform.

Instead of attempting to find a way to continue to use the WISC-R in gifted selection models, it may behoove educators to adopt the use of multiple test instruments, including a nonverbal instrument such as Raven's Progressive Matrices (Raven, 1938) plus a measure of verbal reasoning ability as well as behavioral and motivational indicators. In any case, inclusion/exclusion decisions should never be

based on a single test score, just as no clinical decision should be solely based on any one score or even on one test.

The studies reported here are limited by the same factors inherent in any quasi-experimental design as well as by the limitations of archival and cross-sectional research. Attempts were made to control for gender and to examine ethnic and risk level effects. It should be pointed out, however, that the original sample from which subsamples were drawn was not a random, multivariate normal sample from the entire population. Instead, this was a sample of children referred by parents, teachers, and central nomination for assessment because each had in some way demonstrated the potential for intellectual giftedness. Almost certainly biases were inherent in that referral process. One of those biases can be seen in the unequal proportions of children from different ethnic backgrounds. Another concerns the under-representation of WISC-R scores from Hispanic children who have English as a second language, and are often tested with other assessment devices.

Self-report questionnaires are a source of error from both under-reporting and over-reporting. Ethnic background was deduced by response of the primary caregiver to a school district questionnaire, and incidence of risk was gathered from information questionnaires provided by teachers and by parents. For example, a child who is 10% Native American may be reported as Native American, while one who is Hispanic/Caucasian may be reported as Caucasian because of beliefs the parents hold about the implications for their child of certain ethnic designations within the education system. The risk factors examined in this work are almost certainly an under-representation of population incidence. Those at risk may have been under-reported both by teachers who have have less contact with particular groups of parents, and by the parents who are overwhelmed by the same environmental stressors that affect their children. It is likely that the more seriously economically and environmentally disadvantaged have less access to health care and may also mistrust an educational establishment that doesn't seem to be addressing their most pressing needs. On the other hand, affluent parents may over-report certain risk factors, such as health problems and emotional problems.

There could be other areas of risk not included in the risk factor questionnaires used by this district. For example, acculturation issues are complex and are not well investigated in these questionnaires. Other than the self-report of cultural differences in the home (in the student-parent questionnaire, Appendix C), no attempt could be made to divide groups on the basis of acculturation, since we did not have access to detailed structured interviews. Lastly, ethnic categories were broad and included diverse groups within some single categories. For example, the one category "Hispanic" included Latinos, Cubans, Puerto Ricans, and Hispanics. The group "Asians" included children of Japanese and Chinese background. It may be that more differences exist within these heterogeneous groups than across our ethnic categories.

In terms of the test data itself, some subtest scores were frequently missing from the data since the school district, due to time and financial constraints, does not routinely administer all of the subtests of the WISC-R. Therefore Coding, Mazes, Comprehension, and Digit Span could not be included in the multivariate modeling phase of the study. It may be possible to find more accurate selection models with the WISC-R if those subtests are included.

Finally, the sample was drawn entirely from the San Diego area. Results may not generalize to other geographic areas: San Diego is a metropolitan area with a preponderance of Latinos in its Hispanic population and a significant proportion of first and second generation Asians in its Asian population. As noted above, the majority of the students in the district are non-Caucasian.

On the other hand, the sample reported here is derived from an ethnically diverse district that has consistently shown a commitment to identification of disadvantaged students for gifted programs. The sample does consist of the entire population of children referred and subsequently administered the WISC-R as part of the selection process for gifted education programs. The number of non-Caucasian children, particularly in the gifted base rate sample, is larger and more diverse than any previously reported. Moreover, a completely balanced sample of 1438 was randomly selected from an overall

sample of 8396. Lastly, if risk factors were indeed under-reported and ethnic groups were heterogeneous, the likelihood of finding clear differences between groups would be decreased. Nevertheless, differences were found.

This work has demonstrated clear differences in pattern of abilities across ethnic backgrounds and across levels of risk in the children's environment. Population rates of VIQ-PIQ discrepancies have been documented, and the importance of the difference between statistical significance and clinical rarity across ethnic groups has been illustrated. No single model using the WISC-R was found to provide proportionately equal access to gifted programs. Individual models based on ethnic background failed to achieve ethnic balance, since individual models over-selected African-Americans, Asians, and Caucasians relative to Filipinos and Hispanics. Therefore, **use of the WISC-R in a diverse population to select a balanced group was demonstrated to be inappropriate.**

Other instruments, such as Raven's Progressive Matrices, need to be tested in such a large, multicultural population, as has been repeatedly recommended (Baska, 1986; Pearce, 1983; Valencia, 1984). Some combination of assessment devices that account for motivation as well as intellectual potential may need to be evaluated. Given that we find a way to identify greater proportions of disadvantaged children with high potential, the focus then must turn to finding ways to ensure that these "different" children express that potential. The children we identify with alternative methods may not be the verbally gifted, behaviorally compliant children who currently populate the gifted classrooms in this district. Further work will need to focus on changes in the enrichment programs themselves, to enable teachers of the gifted to unlock and direct the potential these children demonstrate. Improved identification is certainly a goal that needs to be met, but it will be an empty victory if it is achieved and the children so identified fail to be able to express that potential in ways that add to their own growth as well as the growth of their cultures and societies.

There is a dearth of data on how best to nurture particular kinds of talents: that lack of research-based knowledge, combined with administrative inflexibility in the use of resources (particularly in a climate of increasing budget constraints), bodes poorly for children whose giftedness is expressed not so much in verbal domains as in other intellectual areas. We have seen from the experience at Berkeley and other universities (D'Souza, 1991) the disturbing outcome of including individuals who would not have qualified based on established, traditional, uniform criteria into a system which does not change to fit their needs. Dropout rates are high, and we do not know the long-term negative effects of the experience for those who are not able to complete the program. Future studies are needed to examine ways to develop effective educational programs for diverse classes of very young gifted children, and longitudinal studies will be necessary to evaluate the effectiveness of the interventions so developed. Understanding of motivational principles, group process, and cultural as well as individual differences in achievement needs and in pattern of abilities is vital for the development of interventions to provide mastery experiences for these children early in the educational process. Only then can we provide more effective strategies for engaging these children in the life-long growth and development of the potential they as individuals possess.

Appendix A.

San Diego City Schools
Educational Services Division
Gifted and Talented Education

Teacher Nomination Form

Date_____

Student Name_____ Birth Date_____ Sex____ Ethnic Code____
(last) (first) (mi)

School_____ Grade_____ Room Number_____

GUIDE FOR IDENTIFICATION OF PROSPECTIVE GATE CHILDREN

Please rate (name)_____ on each of the following characteristics. This is a five-point scale with the lower end of the scale (#1) indicating lower than average performance and the upper end (#5) indicating excellent or exemplary performance.

	1	2	3	4	5
I. PERSONAL					
1. Curious; asks many questions	___	___	___	___	___
2. Self-motivated; requires little external direction or encouragement	___	___	___	___	___
3. Likes to organize people and structure activities	___	___	___	___	___
4. Generates many ideas, questions, and suggestions	___	___	___	___	___
5. Flexible; adapts readily to new situations	___	___	___	___	___
6. Impatient with routine tasks	___	___	___	___	___
II. EXPRESSION					
7. Vocabulary beyond chronological age or grade level	___	___	___	___	___
8. Advanced skill in written expression	___	___	___	___	___
9. Proficiency in oral expression	___	___	___	___	___
III. THOUGHT PROCESSES					
10. Quick and accurate recall of factual information	___	___	___	___	___
11. A storehouse of information on a variety of topics	___	___	___	___	___
12. Readily recalls visual information	___	___	___	___	___
13. Readily recalls auditory information	___	___	___	___	___
14. Generalizes learning from one experience to another	___	___	___	___	___
15. Finds differences and similarities in events	___	___	___	___	___
16. Understands concepts without extensive concrete examples	___	___	___	___	___
17. Can establish relationships between seemingly unrelated concepts and ideas	___	___	___	___	___
18. Is insightful about cause and effect relationships	___	___	___	___	___

IV. PRODUCTION AND OUTPUT	1	2	3	4	5
19. Displays a great deal of imagination	___	___	___	___	___
20. Manipulates ideas (i.e., makes changes and elaborates upon them)	___	___	___	___	___
21. Concerned with improving or adapting objects and systems	___	___	___	___	___
22. Capable of intense concentration on tasks of interest to her/him	___	___	___	___	___
23. Does not give up easily when confronted with a challenge; shows determination in achieving goals	___	___	___	___	___
24. Offers unique, clever responses to questions	___	___	___	___	___
25. Resourceful, knows where to find answers	___	___	___	___	___
V. ACHIEVEMENT					
26. High performance (grades) in a particular subject, e.g., math, language arts, science, other_____	___	___	___	___	___
27. Achieves at a high educational level	___	___	___	___	___
VI. LEADERSHIP					
28. Has strong communication skills; gets ideas across effectively	___	___	___	___	___
29. Assumes leadership role easily	___	___	___	___	___
30. Facilitates and directs efforts	___	___	___	___	___
VII. OTHER CHARACTERISTICS					
31. Dominates situations	___	___	___	___	___
32. Expressive of thoughts and opinions	___	___	___	___	___
33. Compulsive about work and work habits; strives for perfection	___	___	___	___	___
34. Becomes involved in task, loses awareness of time	___	___	___	___	___
35. Persistent in pursuing discussion beyond cutoff point	___	___	___	___	___
36. Appears inattentive, withdrawn (daydreams)	___	___	___	___	___

Prepared by_____ Recommended? Yes____ No____
(Teacher)

Reviewed by_____ Recommended? Yes____No____
(Administrator/Designee)

Appendix B.

San Diego City Schools
Educational Services Division
Gifted and Talented Education

TEACHER NOMINATION FORM

Date_____

Name_____ Birth Date_____ Sex_____Ethnic Code_____

School_____ Grade____Track _____Room Number_____

SOCIAL/ENVIRONMENTAL VARIABLES

Please check all items that apply:

1. ENVIRONMENTAL
   - _____ Lacks preschool/kindergarten experience
   - _____ Irregular attendance
   - _____ Transiency (3 or more school moves)
   - _____ Limited home enrichment opportunities (availability of books, periodicals, family interaction, family outings)
   - _____ Home conflicts:
     - _____ Responsibilities and study time
     - _____ Excessive child care responsibility
     - _____ Working to help support family
     - _____ Overcrowding — no study area
     - _____ Inconsistencies in the home
2. ECONOMIC
   - _____ Economic hardship
   - _____ Single parent head of household
   - _____ Unemployment
3. LANGUAGE
   - _____ Primary language of parent and/or student is other than English
   - _____ Not proficient/fluent in English
   - _____ Uses non-standard English
   - _____ Student enrolled in Second Language Immersion Magnet (SLIM)
4. CULTURAL
   - _____ Limited home/school communication
   - _____ Experience in dominant culture is limited
   - _____ Cultural values and beliefs differ from dominant culture

5. SOCIAL/EMOTIONAL

_____ Child abuse: physical_____ mental_____ neglect_____

_____ Emotional/adjustment problems
- _____ Working with district counselor
- _____ Working with social worker
- _____ Utilizing psychological services
- _____ Other:_____

_____ Significant home factors
- _____ Separation
- _____ Divorce
- _____ Death

_____ Extended absence of parent
- _____ Military
- _____ Employment
- _____ Other:_____

_____ Family
- _____ Single parent
- _____ Remarriage/step-parent

6. HEALTH

_____ Designated instructional services
- _____ PHDIS
- _____ Speech and language
- _____ Vision
- _____ Hearing
- _____ Adaptive P.E.

_____ Severe allergies

_____ Asthma

_____ Frequent medical/health referral

_____ Regularly prescribed medication

_____ Other:_____

Prepared by_____ Recommended? Yes_____ No_____
(Teacher)

Reviewed by_____ Recommended? Yes_____ No_____

Appendix C

San Diego City Schools
School Services Division
Gifted and Talented Education

STUDENT/PARENT INFORMATION FORM

Student Name: Date_____

_____Birth Date_____Sex___School_____
(Last) (First) (mi)

Address_____Mother's name_____Occupation_____
(Street) Work Phone_____

_____Father's name_____Occupation_____
(City) (State) (Zip) Work Phone_____

Grade_____ Room Number_____ Track_____ Home Phone_____

Schools Attended Grade Dates Attended

_____

_____

_____

1. Names and ages of brothers and sisters:_____

_____

2. Describe your child's attitude toward school:_____

_____

3. List any special interests, talents, and skills your child may have:_____

_____

4. What special lessons, training or learning opportunities has your child had outside of school?

_____

_____

5. To help us know more about your child, please check any of the following that apply:

- ❑ allergies
- ❑ asthma
- ❑ frequent absences
- ❑ prescribed medications
- ❑ parent in military
- ❑ frequent parent absence
- ❑ parents separated
- ❑ single parent
- ❑ remarriage/step-parent
- ❑ recent death/significant illness in family
- ❑ 3 or more schools attended
- ❑ no kindergarten or pre-school experience
- ❑ additional language(s) spoken in home
  List:_____

6. Has your child been previously assessed? ❑ yes ❑ no If yes, when?_____

7. What other things would you like us to know that would assist us in assessing your child?

_____

_____

Name of person completing this form_____ Relationship to student_____